STAT

MEMORANDUM FOR:

Neil,

Believe Helen recognizes our concern on what comes next after her job by the last sentence on page 2 of her memo. (N.B. The COINS Eval Panel is her Panel in this case.)

11 Sep 68
(DATE)

Rudy

FORM NO. 101   REPLACES FORM 10-101
1 AUG 54   WHICH MAY BE USED.

1 April 1968

EVALUATION OF COINS I

The purpose of this brief paper is to further the process
of clarifying the meaning and the purpose of "evaluation" with
respect to COINS I. Ever since my appointment as chairman of
the COINS Evaluation Panel, I have been bothered by the semantic
ambiguity in the terms "experiment" and "evaluation" as applied
to COINS, and I believe it would be helpful to all concerned with
COINS if we could agree on what is to be understood when we use
these terms.

It has seemed clear to me from the COINS documents that
"experiment" has been used all along in the broader sense of a
"trial, " and not in the narrower sense of a scientific or controlled
experiment. The community has developed and will be trying out
a system that allows analysts to access information in selected
files in other agencies. I suspect that there would be little disagreement
with this understanding of the nature of the COINS "experiment. "

The term "evaluation, " however, causes more difficulties.
I would argue that "evaluation, " like "experiment, " should be
understood in the broader sense of "learning as much as we can
from the COINS experiment, or trial, " in order to determine
whether we are on the right track with this approach to a community
system and to decide on the most promising next steps; it should
not be understood in the narrower sense of measuring the performance
or use of the system against some standard, or base line, or set
of evaluative criteria. I shall try to explain why I believe former
interpretation of "evaluation "is the more realistic one.

Since COINS was not planned as a controlled experiment, the
groundwork has not been laid for comparison in any quantitative sense
of the effectiveness of COINS with whatever informal arrangements
for exchange of information constituted the community "system"
before COINS. In other words, comparison of the "before" and
the "after" is not feasible because we have no data on the "before, "
-- no base line against which to measure the changes brought
about by the introduction of the new system. We cannot "evaluate"
performance and use data on COINS against criteria that must be

met, because no such criteria have been established; nor do we
have any statement of precise requirements or design objectives
that would constitute a sound basis for developing such criteria
now. And we cannot compare the effectiveness of COINS with
any alternative new means of accomplishing the same objective,
because only one of the alternatives was implemented. I am
not implying that things should have been done that were not done,
for I think that in many instances a good case can be made for
introducing a new system in order to give the users experience
with it and to learn how they use and how it might be made to
serve them better.

What we can do in the present circumstance is to learn
as much as possible about the community's experience in attempt-
ing to use COINS and about the technical performance of the
system; and this is the objective of the plan the COINS Evaluation
Panel was asked to develop -- namely, a plan for acquiring,
analyzing and presenting data on system performance, system
utilization, and user reaction.

The data on the performance on the COINS hardware and
software will be quantitative in nature and should be useful in
indicating what needs to be done to make the system perform better
and to overcome user dissatisfaction caused by inadequacies in
technical performance. I would say that such data will give us
performance measurements, but not evaluation in the strict sense
of meeting certain standards or criteria.

The more significant question of potential value or utility
of such a system to the community and decisions as to where we
go from COINS I will depend more on the response of the users.
I believe the main goal of our analysis of utilization and user
reaction should be understanding of the factors that account for
differing levels of use, users' reactions, and their views as to
the potential utility of such a system. I would expect that conclusions
as to what has been learned from the COINS experiment and decisions
as to the next steps will have to be products of informed judgment --
informed, I would hope, by the data we plan to present. In this
connection, it would be helpful to the COINS Evaluation Panel to
understand more clearly what kinds of decisions must be made,
and when, on the basis of the data to be presented.

In the last analysis, use is the best indicator of the value, or utility, of an information system, but as of now there is no basis for saying how much use must occur before the COINS system or any file within it has demonstrated its utility. Statistics on use must be treated with considerable care. Even if a new kind of information system has reached the operational stage, and COINS has not, it still may take quite a long while for the community of users it serves to learn to make effective use of it and to adjust their work habits to it. We shall be proposing, therefore, to gather and analyze not only quantitative data on use but also detailed information from selected users on the reasons for their actions and views and descriptive information on the differing conditions in the participating agencies that may help to explain different patterns of use.

We shall propose to gather and analyze data on both performance and use for a full year, because we understand that the system will continue to evolve. Performance data may change somewhat as equipment changes are made and as software is improved in order to achieve better performance. Use patterns should change as users learn more about the system. Changes in both performance and use data should occur if new files are added, or if files are removed from the system. COINS has to be viewed, therefore, as a dynamic system, and changes in it should be reflected in the data presented.

The COINS Evaluation Panel was also asked to recommend other data gathering and analysis activities that would be desirable in connection with system evaluation. We have not yet turned our attention to this task, but we may well give further thought to the fourth point made in our response to the COINS files question. In the case of any file recommended for inclusion in COINS, it may be desirable to require that data be presented on the extent of present file activity, on interagency communication with respect to such data, and on unmet demands for such data throughout the community. There would then be a base line for comparison with community use of the file after it becomes available through COINS.